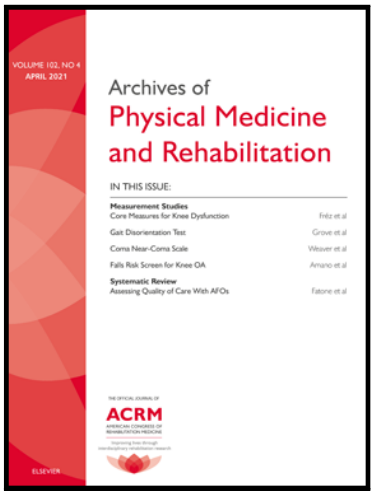
2021-04-23

Is aerobic or resistance training the most effective exercise modality for improving lower extremity physical function and perceived fatigue in people with multiple sclerosis? A systematic review and meta-analysis

Taul-Madsen, L

10.1016/j.apmr.2021.03.026
Archives of Physical Medicine and Rehabilitation
Elsevier

Journal Pre-proof



Is aerobic or resistance training the most effective exercise modality for improving lower extremity physical function and perceived fatigue in people with multiple sclerosis? A systematic review and meta-analysis.

Laurits Taul-Madsen MSc. , Luke Connolly PhD ,
Rachel Dennett BSc , Jenny Freeman PhD , Ulrik Dalgas PhD ,
Lars G. Hvid PhD

PII: S0003-9993(21)00307-5
DOI: https://doi.org/10.1016/j.apmr.2021.03.026
Reference: YAPMR 58199

To appear in: *Archives of Physical Medicine and Rehabilitation*

Received date: 28 January 2021
Revised date: 10 March 2021
Accepted date: 25 March 2021

Please cite this article as: Laurits Taul-Madsen MSc. , Luke Connolly PhD , Rachel Dennett BSc , Jenny Freeman PhD , Ulrik Dalgas PhD , Lars G. Hvid PhD , Is aerobic or resistance training the most effective exercise modality for improving lower extremity physical function and perceived fatigue in people with multiple sclerosis? A systematic review and meta-analysis., *Archives of Physical Medicine and Rehabilitation* (2021), doi: https://doi.org/10.1016/j.apmr.2021.03.026

Running head: Exercise modalities in multiple sclerosis

Is aerobic or resistance training the most effective exercise modality for improving lower extremity physical function and perceived fatigue in people with multiple sclerosis? A systematic review and meta-analysis.

Laurits Taul-Madsen[1], MSc.

Luke Connolly[2], PhD

Rachel Dennett[2], BSc

Jenny Freeman[2], PhD

Ulrik Dalgas[1], PhD

Lars G. Hvid[1], PhD

1. Exercise Biology, Department of Public Health, Aarhus University, Denmark

2. School of Health Professions, Faculty of Health, University of Plymouth, UK.

Corresponding author:
Laurits Taul-Madsen
Dalgas Avenue 4
8000 Aarhus C
Denmark
ltm@ph.au.dk
+45 61656063

Objective: The purpose of this systematic review was to investigate whether aerobic training (AT) or resistance training (RT) is most effective in terms of improving lower limb physical function and perceived fatigue in persons with multiple sclerosis (pwMS).

Data sources: Nine databases (MEDLINE, EMBASE, CINAHL, AMED, PEDro, SPORTdiscus, PsycINFO, Web of Science and SCOPUS) were electronically searched in April 2020.

Study Selection: Included studies were randomized controlled trials (RCTs) involving pwMS attending one of two exercise interventions; AT or RT. Studies had to include at least one objective or self-reported outcome of lower extremity physical function and/or perceived fatigue.

Data Extraction: Data was extracted using a customized spreadsheet, which included detailed information on patient characteristics, interventions and outcomes. The methodological quality of the included studies was independently assessed by two reviewers using the TESTEX rating scale.

Data synthesis: Twenty-seven papers reporting data from 22 RCTS (AT=14, RT=8) including 966 pwMS. The two modalities were found to be equally effective in terms of improving short walk test (AT: ES=0.33 [-1.49: 2.06]; RT: ES=0.27 [0.07: 0.47]) and long walk test performance (AT: ES=0.37 [-0.04: 0.78]; RT: ES=0.36 [-0.35: 1.08]), as well as in reducing perceived fatigue (AT: ES=-0.61 [-1.10:-0.11]; RT: ES=-0.41 [-0.80: -0.02]). Findings on other functional mobility tests along with self-reported walking performance were sparse and inconclusive.

Conclusions: AT and RT appear equally highly effective in terms of improving lower extremity physical function and perceived fatigue in pwMS. Clinicians can thus use either modality to target impairments in these outcomes. In a future perspective, head-to-head exercise modality studies are warranted. Future MS exercise studies are further encouraged to adapt a consensus 'core battery' of physical function tests to facilitate a detailed comparison of results across modalities.

Keywords:
Rehabilitation; Exercise; Multiple Sclerosis; Systematic Review.

Abbreviations:

MS: Multiple Sclerosis

pwMS: Person with multiple sclerosis

RCT: Randomized controlled trial

AT: Aerobic training

RT: Resistance training

RM: Repetition maximum

HR: Heart rate

Testex: Tool for assessment of study quality for reporting on exercise

RPE: ratings of perceived exertion

ES: Effect size

CI: Confidence interval

EDSS: Expanded disability status scale

6MWT: Six minute walk test

MSWS: 12-item multiple sclerosis walking scale

SSST: Six spot step test

FSS: Fatigue severity scale

Introduction

Multiple sclerosis (MS) is a chronic, autoimmune, and inflammatory disease of the central nervous system, exemplified through demyelination and axonal loss[1]. As a consequence, multiple symptoms can appear[1-3], with fatigue and walking limitations reported to be among the most debilitating[4-7]. Moreover, an estimated 50% of persons with MS (pwMS) will require a walking aid within 15-25 years after disease onset[8,9]. Since physical function is associated with lowered quality of life at the individual level along with a greater economic burden at a health service and societal level [10,11], it is crucial to diminish progression of disability[12].

While pharmacological treatments appear to have limited beneficial effect on fatigue and walking limitations[13], exercise has proven to be a potent non-pharmacological treatment option, being both safe and eliciting numerous beneficial effects in pwMS [14,15]. Specifically, exercise is an effective way of reducing fatigue[16,17] and improving walking performance [18,19], with the latter often considered to be clinically meaningful[20,21].

Exercise constitutes a number of different modalities known to elicit different physiological adaptations (such as neuromuscular function or cardiovascular function) that in most cases are paralleled by (and perhaps even translated into) improved physical function[22]. A recent review investigating randomized controlled trials (RCTs) of exercise interventions in pwMS reported that the two most applied exercise modalities were aerobic training (AT) and resistance training (RT)[23]. Several studies have reported positive effects of both AT[24-26] and RT[27,28] on parameters directly related to lower extremity physical function (e.g. walking performance, chair rise, stair negotiation) as well as on parameters indirectly related to lower extremity physical function, such as perceived fatigue. However, based on the existing literature it currently remains unknown which of these two common exercise modalities is the most effective in terms of improving physical function and perceived fatigue in pwMS. Despite the somewhat impossible task of matching AT and RT on traditional exercise parameters such as duration, frequency, and intensity, understanding the specific effectiveness of the two different exercise modalities is an important factor for consideration in optimizing exercise prescription in pwMS.

Therefore, the objectives of this systematic review were to investigate which of the two exercise modalities (AT or RT) are the most effective in terms of improving lower extremity physical function and reducing perceived fatigue in pwMS.

Methods:

The present systematic review follows the Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) guidelines on systematic reviews of RCTs[29]. Search strategy, study selection, eligibility criteria, methodology assessment, data extraction and analysis were performed in accordance with a protocol pre-registered in PROSPERO (CRD42020189855).

Definitions:

In this review the following definitions were applied:

Exercise: A form of physical activity that is planned, structured and repetitive, and is undertaken with the objective of improving or maintaining at least one aspect of physical fitness, comprising strength, flexibility or aerobic endurance[30].

Physical activity: Any bodily movement produced by skeletal muscles that requires energy expenditure above resting levels[30].

Physical function: The ability of an individual to perform physical activities of daily living. For the purposes of this systematic review, this particularly relates to lower extremity tasks (e.g. simple/complex/endurance walking, chair rise, stair negotiation) [31].

Perceived fatigue: Subjective sensations of weariness, increasing sense of effort, mismatch between effort expended and actual performance or exhaustion[32].

Resistance training: Performed with external resistance of varying degrees relative to maximal strength provided by either free weights, machines, bodyweight, or some other implements (e.g., resistance bands), either with single or multiple sets of repetitions which may or may not be performed to momentary failure (but are often performed to a relatively high effort)[33].

Aerobic training: Performed using locomotor or ergometer tasks (e.g., walking, jogging, running, cycling, rowing, etc.) in a continuous or intermittent fashion with respect to duration at submaximal intensities of effort, commonly determined relative to either maximal heart rate, heart rate reserve, VO_{2max}, or sometimes using ratings of perceived effort scales[33].

Exercise intensity: For AT, exercise ≤ 63% of Heartrate max (HRmax) was defined as low intensity, 64-76% of HRmax as moderate intensity, and ≥ 77 % of HRmax as high intensity[34]. For RT, exercise ≥ 16 Repetition Maximum (RM) was defined as low intensity (≤ 64% of 1RM), 9-15 RM as moderate intensity (65-79% of 1RM) and ≤ 8 RM as high intensity (≥ 80% of 1RM)[35,36]

Searches

An original search was carried out as part of another review by the same authors in 2018, having the aim to summarize reported adherence and drop-out data from RCT studies of exercise interventions in pwMS.[23]

This search was updated in April 2020. Furthermore, in March 2020, The World Health Organization's International Clinical Trials Registry Platform (ICTRP) http://apps.who.int/trialsearch/, which comprise the 16 primary registries of the WHO registry network and ClinicalTrials.gov, was searched for relevant ongoing trials investigating a head-to-head comparison of AT and RT in pwMS.

Data sources and search strategy

In brief, the search strategy was based on the key terms "multiple sclerosis" OR MS AND exercise OR "physical activity". For full search strategy please see Dennett et al. 2020[23].

The original search was carried out in October 2018 and updated in April 2020.

Two reviewers (LM and RD) conducted the original search in the electronic databases MEDLINE, EMBASE, CINAHL, AMED, PEDro, SPORTdiscus, PsycINFO, Web of Science and SCOPUS limited to scientific research papers being published between January 1993 and October 2018. The same databases were searched from September 2018 to March 2020 by two reviewers (LM and LC) in April 2020. All searches were supplemented by hand searches of reference lists.

Study selection

The following PICO (population, intervention, comparison, outcome) question guided the search and inclusion strategy. "Which exercise modality, AT or RT, is most effective in improving physical function (specifically lower extremity tasks such as simple/complex/endurance walking, chair rise, stair negotiation) and perceived fatigue in pwMS?"

Eligibility criteria

RCT studies involving adults over the age of 18 with a definite diagnosis of MS, regardless of gender, disease duration, MS phenotype or level of disability were considered eligible for inclusion. While all identified studies could be included regardless of location, group/ individual structure, level of supervision, intervention duration, session duration, intensity, progression, frequency, the content had to be either AT or RT; with or without a follow-up period.

Control interventions had to include non-training controls only or active control conditions having no expected effects on the cardiovascular system or the musculoskeletal system, for example stretching were accepted.

Studies had to include at least one objective or self-reported measure of lower extremity physical function (such as simple/complex/endurance walking, chair rise, stair negotiation) and/or perceived fatigue. If reported, measures of cardiovascular function (i.e. maximal oxygen uptake) and neuromuscular function (i.e. maximal muscle strength or muscle power) were also extracted, as these outcomes could, (1) help verify the effectiveness of interventions, and (2) are likely mediators of adaptations in lower extremity physical function.

Data management and selection process

The original search resulted in 93 papers included in the previous review, all of which were considered for inclusion in the present review (see figure 1).

Results from the updated search were exported to EndNote, where duplicates were removed. The remaining papers were imported into Rayyan data management system (rayyan.qcri.org) where titles and abstracts were independently screened for eligibility by two reviewers (LC and LTM). If papers were included at this stage, a full-text reading by the same two reviewers was performed, and any discrepancies were discussed with a third party (LGH). Reasons for excluding full text RCTs were recorded.

Data extraction

Data was extracted using the same spreadsheet as the previous review[23], which included detailed information on participant characteristics (age, gender, disease duration, MS phenotype, disability level, and fatigue as a symptom); modality of the intervention (setting, group/individual structure, level of supervision, intervention duration, session duration, intensity, frequency); content of the intervention (aerobic or resistance); report of adverse events, % drop-out, and adherence during the intervention period and at any follow up. Furthermore, an additional customized spreadsheet was made to extract information on all outcomes of lower extremity physical function, perceived fatigue and measures of cardiovascular and neuromuscular function. Data extraction was completed by two reviewers (LC and LTM).

Quality assessment

The methodological quality of the included studies was independently assessed by two reviewers (LTM and LC) using the 'Tool for assessment of study quality for reporting on exercise' (TESTEX) rating scale[37]. Any discrepancies were discussed and resolved between the two reviewers.

Synthesis of results

In addition to the qualitative analysis (summary of identified studies and their data), we also performed quantitative analysis by calculating sample-size weighted averages across selected studies. A minimum of two studies was required in order to conduct a meta-analysis. Random effects meta-analyses comprising data on physiological adaptations, short walking tests, long walking tests and perceptions of fatigue were conducted by using *Meta-Essentials version 1.5* designed for Excel.[38] Intervention effect sizes (ES) (between-group differences) for different outcomes at post-treatment, were calculated using Hedges' g statistic, along with 95% confidence intervals (CIs) around the estimated effect-size. Also, if data was available and adequate, we performed a weighted regression of all study ES as a function of intervention duration and frequency (weeks and number of sessions) as well as intervention intensity, as these factors were hypothesized to impact the outcomes[39]. Of note, this approach was done to establish specific within-modality information only. ES were interpreted as follows: small = 0.14, moderate = 0.31, large = 0.61 based on empirical data from 99 meta-analyses examining the effects of rehabilitation/exercise[41]. Statistical heterogeneity was quantified using Higgins' I^2 statistic, and was interpreted as follows: heterogeneity: > 50%, no or limited heterogeneity: < 50%[42].

If studies reported on more than one outcome in each domain (e.g. physiological adaptations such as knee extensor and knee flexor muscle strength as well as perceptions of fatigue using different questionnaires), an average was calculated and used for the meta-analyses.

[INSERT FIGURE 1]

Results

Study characteristics

As depicted in figure 1, the search yielded 2117 hits. After removal of duplicates, 1538 papers remained for the screening process, with 12 of these assessed for full-text reading. Five papers were included, which with the addition of 22 papers from the previous review, resulted in a total of 27 papers being included in the qualitative and quantitative synthesis.

The 27 papers reported 22 RCTs (AT (n=14), RT (n=8)) which involved a total of 966 pwMS. As seen in table 1, Expanded disability status scale (EDSS) ranged from 1.5-7 while disease duration ranged from 2.7-18.6 years. The duration of AT interventions ranged from 3-26 weeks (involving 9-48 sessions) with the intensity being deemed moderate (n=5)[43-47], high (n=4)[26,48-50], or unknown (no information, n=5)[25,51-54]. The duration of RT interventions ranged from 8-24 weeks (involving 15-48 sessions) with the intensity being deemed moderate (n=1)[55], high (n=4)[28,56-58], or unknown (no information, n=3)[40,59,60]. Due to the missing information and the use of divergent scales of exercise intensity for both AT (e.g. % of HR_{max}, RPE, % of $VO2_{max}$, % of Peak Power) and RT (% of 1RM, % of bodyweight, absolute weights), we were unable to perform weighted (moderator) analysis using this parameter. Two[25,60] of the 22 identified RCTs reported a primary outcome that was not based on a sample size calculation. Ten papers[26,28,44,48-50,52,54,55,57] of the 22 identified RCTs reported a primary outcome based on a sample size calculation, with five of these having a primary outcome aligned with the purpose of the present systematic review.

The median TESTEX score of the included studies was nine out of 15. Detailed information on the scores can be found in Table 2.

[INSERT TABLE 1]

[INSERT TABLE 2]

[INSERT TABLE 3]

Physiological adaptations

Seven of the 14 AT studies[26,45-48,50,52] reported a between-group change in aerobic capacity, with four of these[45,46,48,52] reporting a statistically significant improvement (Table 3). The meta-analysis showed an overall large effect on aerobic capacity, ES=0.88 [0.25: 1.50], p=0.001, I^2=78% (Figure 2). Aerobic capacity ES was not positively associated with AT intervention duration (weeks: slope -0.03, r^2=0.06, p=0.563; number of sessions: slope 0.00, r^2=0.00, p=0.97.

In regard to RT studies, seven out of nine studies[28,40,55-59] reported a between-group change in one or more strength measurements, with five of these changes being reported as statistically significant. The meta-analysis showed an overall large effect of RT on muscle strength, ES=0.86 [0.02: 1.70], p=0.013, I^2=75% (Figure 2). Strength ES appeared to be positively associated with RT intervention duration (weeks: slope 0.08, r^2=0.25, p=0.104; number of sessions slope 0.06, r^2=0.44, p=0.019).

[INSERT FIGURE 2]

Performance on short walking tests

Three out of the 14 AT studies[43,50,52] reported a between-group change in short walking tests, with one of these changes[43] being reported as statistically significant (Table 3). An overall moderate effect was observed in the meta-analysis, ES=0.33 [-1.49: 2.06], p=0.20, I^2=69% (Figure 3). Short walk ES was not positively associated with AT intervention duration (weeks: slope -0.32, r^2=1.00, p=0.011; number of sessions: slope -0.10, r^2=0.68, p=0.15).

Six RT studies[28,40,56-59] reported a between-group change in any short walking test, with one of these reporting a significant change (Table 3). The meta-analysis showed a moderate effect of RT on short walking performance, ES=0.27 [0.07: 0.47], p=0.006, I^2=0% (Figure 3). Short walk ES was not positively associated with RT intervention duration (weeks: slope -0.02, r^2=0.64, p=0.51; number of sessions slope -0.01, r^2=0.42, p=0.59).

[INSERT FIGURE 3]

Performance on long walking tests

Of the long walking tests, the Six minute walk test (6MWT) was the most used in AT studies. Five[44,48-50,52] out of the seven[25,43,44,48-50,52] studies investigating performance on a long walking test used this test. The meta-analysis showed an overall moderate effect of AT on the performance during long walking tests, ES=0.37 [-0.04: 0.78], p=0.026, I^2=43% (Figure 4). Long walk ES was not positively associated with AT intervention duration (weeks: slope 0.01, r^2=0.03, p=0.70; number of sessions: slope 0.01, r^2=0.14, p=0.36).

Four RT studies[28,55,57,58] reported a between-group change in any long term walking test, with one of these reporting a statistically significant finding and the meta-analysis showing a moderate effect of RT on long walking test performance, ES=0.36 [-0.35: 1.08], p=0.11, I^2=48%

(Figure 4). Long walk ES was positively associated with RT intervention duration (weeks: slope 0.07, r^2=0.87, p=0.025; number of sessions slope 0.07, r^2=0.87, p=0.025).

[INSERT FIGURE 4]

Performance on functional mobility tests

Only one[52] of the AT studies investigated effects on the performance of a functional mobility test, and reported a statistically significant change between groups.

Five[28,56-58,60] of the RT studies investigated the performance on a functional mobility test between groups, with two[56,58] of these changes being reported as statistically significant.
As the aim of this present review was to evaluate differences between modalities, we were not able to conduct a meta-analysis on this outcome.

Self-reported walking performance

Two of the AT[50,52] studies reported a between group change in self-reported walking performance (both 12-item Multiple Sclerosis Walking Scale (MSWS)), with one of these[52] being reported as statistically significant. The meta-analysis of AT on self-reported walking performance showed a negligible effect, ES= -0.04 [-2.34; 2.26}, p=0.82, I^2=0% (Figure 5).

Of the RT studies, two[57,58] reported a between group change in self-reported walking performance (both MSWS), with one of these[58] being reported as statistically significant. The meta-analysis of RT on self-reported walking performance showed a negligible effect, ES= 0.07 [-5.20; 5.33}, p=0.88, I^2=66% (Figure 5).

[INSERT FIGURE 5]

Perceptions of fatigue

Nine of the 14 AT studies[26,43-45,47,51-54] reported a between-group change in any measure of perceived fatigue, with four[26,43-45] being reported as statistically significant. The meta-analysis showed a large effect of AT on perceptions of fatigue, ES=-0.61 [-1.10:-0.11], p=0.005, I^2=58% (Figure 6). Improvements in perceived fatigue ES was not positively associated with AT intervention duration (weeks: slope -0.05, r^2=0.00, p=0.85; number of sessions: slope 0.03, r^2=0.31, p=0.052).

Of the RT studies, three[55,57,61] reported a between-group change in any measurement of perceived fatigue, with all of these changes being reported as statistically significant. The meta-analysis of RT on perceived fatigue showed a moderate effect, ES=-0.41 [-0.80: -0.02], p=0.00, I^2=0% (Figure 6). Improvements in perceived fatigue ES was not positively associated with RT intervention duration (weeks: slope 0.10, r^2=0.38, p=0.63; number of sessions slope 0.05, r^2=0.38, p=0.63).

[INSERT FIGURE 6]

Comparison between modalities

While both interventions were shown to elicit adaptations in favor of exercise, we were not able to detect differences in any outcomes between the two different exercise modalities as evidenced by the comparable effect sizes and overlapping confidence intervals.

Discussion

Based on our findings, AT and RT present themselves as broadly equivalent modalities in terms of improving lower extremity physical function (walking performance) and reducing perceived fatigue, with meta-analyses revealing moderate-large effect sizes. Of note, only 14 out of 23 studies reported physiological adaptations thereby limiting the in-depth understanding of the potential mechanistic effect(s) leading to an improvement in physical function (i.e. the translational potential).

Physiological adaptations

Although only seven out of 14 [26,45-48,50,52] AT studies reported a between-group change in aerobic capacity, the observed large effect size (ES=0.88 [0.25; 1.50]) of AT on aerobic capacity corroborate findings of a previous review[62] (ES=0.63 [0.00; 1.26]) using broader inclusion criteria (e.g. by including small pilot studies). Altogether, these provide clear evidence underlining AT as a highly effective intervention targeting the cardiovascular system in pwMS.

The observed large effect size of RT studies on lower extremity muscle strength (ES=0.86 [0.02; 1.70]) corroborate findings previously reported by Jørgensen et al.[63], who in a systematic review and meta-analysis including isokinetic dynamometry determined muscle strength, reported an ES of 0.45 [0.18; 0.72] following RT.

Overall, the physiological adaptations observed by the present systematic review verify that AT and RT interventions overall work as intended, thereby establishing the potential for a translation into improvements in mobility aspects of lower extremity physical function along with reduction in perceived fatigue.

Physical function - walking tests

The identified AT studies predominantly focused on the longer walk tests, with only three studies[43,50,52] investigating the effect on the short walk tests. Despite the moderate ES on the short walk test (ES=0.33 [-1.49; 2.06]; data presented as walking speed) observed in the present systematic review, CIs indicate a high degree of uncertainty. This corroborates the findings of Pearson et al.[19], who reported an ES=-1.96 [-2.67; -1.25] (data presented as walking time). Of note, both findings are based on very few studies (three in the present systematic review and two in the study by Pearson and colleagues), and should therefore be interpreted cautiously. Participants in two of the three identified studies in the present review were relatively high functioning at baseline, based on their short walk test performance and low EDSS [50,52], potentially leaving little room for improvement (due to a ceiling effect). More studies are needed to establish a robust insight into the effects of AT on short walk tests, ideally by involving pwMS who are ambulatory across a wider range of disability levels, especially in severely disabled pwMS having substantial walking limitations.

Of the seven studies[25,43,44,48-50,52] investigating the effect of AT on the long walk tests, three of these[25,43,49] had a large ES. Yet, the meta-analysis showed an overall moderate ES of AT on this outcome (ES=0.37, [-0.04; 0.78]), that appeared quite certain based on CIs. As for the two aforementioned studies involving relatively high functioning participants at baseline[50,52], their long walk test performance was also quite high (6MWT >575m), again potentially leaving little room for improvement. Following 12 weeks of AT, an ES=-0.14 [-0.62; 0.34] was observed on the 6MWT in Baquet et al.[50] whereas an ES=0.33 [-0.34; 1.01] was observed in the study of Feys et al.[52] Interestingly, participants in the study by Feys et al.[52] performed specific walking/running exercises that may have been more beneficial for performance on the long walk test (moderate ES=0.33) compared to short walk test (negligible ES=0.00). Another study whose intervention involved specific walking exercises, was Dettmers et al.[25] who on maximal walking distance observed a moderate ES=0.47 [-0.25; 1.22].

Of the five studies[28,56-59] investigating the effect of RT on short walk test performance, three studies[28,56,59] detected a moderate ES corresponding to the ES of the meta-analysis (ES=0.27 [0.07; 0.47]).

Previously, the effect of RT on the performance on a short walk test has been summarized in a review[64] and in a meta-analysis based on only one study[19]. However, to our knowledge, this is the first systematic review to perform a meta-analysis on RT studies alone, examining the effects on short walk tests (and walking performance in general).

On the long walk test, four RT studies[28,55,57,58] were included in the meta-analysis which showed a moderate ES (ES=0.36 [-0.35; 1.08]), with CIs displaying some degree of uncertainty. These variable results are in line with previous reports[64]. Of note, Kjølhede et al.[58] was the only study showing a large beneficial effect of RT on long walk test performance, ES=1.07 [0.34; 1.86]. Potentially, this is because of the length of the intervention (24 weeks), compared to the shorter interventions in the other studies (10 weeks[55,57] and 12 weeks[28]). This was supported by our weighted (moderator) regression analysis, showing a positive association between intervention duration (weeks and number of sessions) and ES.

Only a few studies investigated the effect of AT[50,52] or RT[57,58] on self-reported walking performance. Based on the two identified studies in each modality, meta-analyses showed a negligible effect on MSWS (AT, ES=-0.04 and RT, ES=0.07), despite both modalities being effective on all objective walking outcomes. As these results are sparse and somewhat inconclusive, they should be interpreted cautiously. Speculatively, they may indicate that adaptations in objectively measured outcomes precede self-reported outcomes, which is somehow contradictory to what has been shown previously[65], and/or that adaptations in self-reported outcomes are limited due to a potential ceiling effect.

Physical function - functional measurements

While walking performance is an essential aspect of lower extremity physical function, our sparse and inconclusive findings reveal an existing knowledge gap in terms of how the two exercise modalities (AT in particular) might impact other measures such as chair rise, six spot step test (SSST) and stair negotiation. This is problematic, since complex walking tests such as the SSST[66] along with highly physically demanding walking tests such as stair negotiation[67], have the potential to give a more in depth picture of patients walking ability. Such tests incorporate not only acceleration and endurance, but also other components such as coordination and balance which are recognized as being important for general physical function. Hence, future AT as well as RT studies should incorporate such complex functional tests in their test battery.

Fatigue measurements

Nine studies[26,43-45,47,51-54] investigated the effect of AT on perceived fatigue. In the majority of these a moderate-large ES[26,43,45,47,54] was observed, with an overall large ES as determined by our meta-analysis (ES=-0.61, [-1.10; -0.11]). This adds further weight to findings of previous systematic reviews (including a Cochrane review) in this area[17,68], with the combined evidence indicating that AT is effective in reducing perceived fatigue.

In this present systematic review and meta-analysis, only three studies[55,57,61] investigated the effect of RT on perceived fatigue. Hence, whilst remaining cautious in our interpretation, data indicate a moderate and beneficial effect of RT on perceptions of fatigue, ES=-0.41, [-0.80; -0.02]. This provides further evidence for already existing guidelines[16].

Comparison between modalities

We did not detect any apparent differences in the magnitude of effect on physiological adaptations in the two exercise modalities. Many components such as duration, frequency and intensity should be taken into account when comparing the two modalities. The average frequency and duration was somewhat comparable between the two exercise modalities (AT: 3

days/week*11 weeks (range 3-26 weeks), 28 sessions (range 9-48 sessions); RT: 2 days/week*11 weeks (range 8-24 weeks), 25 sessions (range 15-48 sessions)), along with the intensity being moderate-to-high in both AT and RT. A plausible explanation for the lack of association between intervention duration (weeks and number of sessions) and meta-analysis ES is that the majority of interventions had durations of 8-12 weeks involving 16-24 sessions. The only exceptions showing positive associations were for RT on muscle strength and long walk test, respectively, although likely driven by one study only[58] having a much longer intervention duration (24 weeks, 48 sessions) compared to the remaining RT studies. Unfortunately the quantity and quality of the reported exercise intensity data (missing information, use of divergent scales of exercise intensity) did not allow us to examine the associations between exercise intensity and meta-analysis ES within each modality. Since factors such as duration, frequency and intensity are crucial for the extent of adaptations[39], further studies seem warranted to help advance our understanding of any potential dose-response association between general exercise parameters (e.g. duration, frequency and intensity) and physiological as well as functional adaptations in pwMS.

To our knowledge, only one pilot study[69] has previously performed a head-to-head comparison of the two modalities, finding no difference in either lower extremity physical function as measured by the six minute walk test and the timed up and go, or in perceived fatigue measured by the Modified Fatigue Index Scale. However, only n=19 participants finished this cross-over study having an eight week wash-out period. Adaptations from exercise interventions may last as long as 12[24] or 24[58] weeks, hence, one must be cautious when interpreting results from this pilot study[69].

Resembling the observations in physiological adaptions, no difference was observed in the magnitude of change on short or long walking tests with AT or RT. All meta-analyses on the walking tests had comparable moderate ES, although data – based on CIs – appeared most robust for short walk with RT and for long walk with AT, respectively. While this is likely influenced by the number of studies for each meta-analyses, it may also be due to physiological adaptations that are intuitively associated with certain aspects of walking (AT: increment in aerobic capacity associated with walking endurance; RT: increment in muscle strength associated with walking acceleration)[70]. While the present findings are aligned with previously reported findings in systematic reviews and meta-analyses[18,19], these were based on a limited number of RCT studies (as the search was performed March 2014)[19] or a combination of RCT and non-RCT studies, different exercise modalities, and different measures of walking performance (self-reported as well as clinician-rated short and long walking performance)[18]. The novel approach of the present systematic review, apart from updating existing evidence, was to include RCTs only, clearly separate study findings across the two most common exercise modalities, and uphold a clear distinction between the selected walking performance outcome measures.

Both modalities were found to be effective in terms of reducing perceived fatigue, with a large ES observed for AT and a moderate ES for RT. While Andreasen et al.[71] in their systematic review previously reported RT to be slightly more effective than AT in terms of reducing perceived fatigue, Heine et al. [17] in their Cochrane systematic review and meta-analysis reported the opposite (applying a broader definition of exercise modalities). In context of the two exercise modalities and their effect on perceived fatigue, Rooney et al.[72] performed a systematic review and meta-analysis and found a strong association between aerobic capacity and perceived fatigue (r=-0.47 [-0.64;-0.25]), but only a moderate association between muscle strength and perceived fatigue (r=-0.22 [-0.40;-0.03]).

Translational or parallel improvements?

Assessment of physiological adaptations are important due to two aspects. First, it is a simple way of validating exercise efficacy as effects on these basic primary (sensitive) physiological targets are expected (i.e. AT expectedly improve aerobic capacity while RT expectedly improve muscle strength). Second, physiological adaptations may be a prerequisite for improvements in physical function, thereby having a translational effect. Interestingly, the findings from the present systematic review and meta-analyses suggest that improvements in lower extremity physical function can be achieved via different physiological pathways (i.e. cardiovascular system or neuromuscular system). At least, we observed parallel improvements in physiological adaptions and in physical function. However, since only a limited number of studies reported parallel data of both physiological parameters and physical function of the same outcome (see Table 3) and since even fewer studies report associations between changes in these outcomes, we were unable to perform any analysis of association. A small number of studies have reported data supporting an exercise-induced translational link, i.e. between improvements in muscle strength and Fatigue Severity Scale (FSS)[61], aerobic capacity and FSS[46], as well as muscle strength and Timed 25-Foot Walk, two minute walk test, five repetition sit-to-stand and stair climb[58]. This is nevertheless challenged by the fact that lower extremity physical function relies on different physiological systems, and adaptations in just one system may elicit little translational response. Also, in high-functioning pwMS the ceiling effect of many commonly used walking measures may mean that changes in performance are not detectable. Nevertheless, physiological adaptations can still be achieved, building physiological reserve capacity as well as improving general health thereby potentially postponing the onset of future physical functional limitations. In order to advance our understanding of any translational link, more studies examining the association between exercise-induced physiological adaptations and measures of physical function are required in pwMS across the entire disability span. This could also help elucidate why some pwMS have a positive effect of an exercise intervention whereas others do not (i.e. responders vs. non-responders).

Clinical and research implications

The present study findings emphasize the importance of providing structured intensive AT and/or RT when aiming to improve lower extremity physical function (along with physiological adaptations). While many different exercise modalities exist, AT and RT have consistently been shown to be among the most effective in terms of positively affecting numerous different domains[22]. As the two modalities proved somewhat comparable (based on magnitude of ESs), it implies that clinicians could use either modality to target impairments in lower extremity physical function - we suggest patient preference be central to this decision to optimize the likelihood of them sustaining exercise over long term. The inconsistency in reporting across studies, emphasize the need for using a “core battery” of physical function tests, as previously proposed[73]. This would enable comparability of findings across studies and facilitate generation of more robust evidence, which is essential for clinicians' decision-making. Moreover, exercise studies should report data for the physiological outcomes they are targeting. This would advance our understanding of potential translational links between physiology and function. Finally, future studies should compare the modalities directly by performing a head-to-head study to establish whether differences in outcomes exist.

Study limitations

The present systematic review and meta-analyses, provides a detailed and comprehensive overview of the RCTs investigating the effect of AT and RT on lower extremity physical function and perceived fatigue. However, some methodological considerations deserve mentioning. First, the majority of identified studies included patients with mild-moderate disease severity, making the results applicable for this subgroup of patients only. Second, more studies are needed to elucidate effects of AT and RT in pwMS with higher levels of disability, including those who are non-ambulatory (EDSS ≥ 7.0), which is a problem that has been exposed previously[74]. Third, this systematic review provides an overview of existing studies evaluating the two modalities, and hence is not able to provide a direct comparison. To provide such information, a well-considered head-to-head study of the two modalities, designed to diminish the difference in intensity and volume, is needed. Finally, our review focused on either solely AT or RT. As such, we cannot comment on the effectiveness of interventions which combine these two exercise modalities or use other exercise modalities (for example pilates, yoga, balance).

Conclusions

Based on knowledge from existing RCTs, aerobic training (AT) and resistance training (RT) appear comparable in improving lower extremity physical function (walking performance in particular) and perceived fatigue. Although substantial physiological adaptations were observed, conclusions about the underlying mechanisms for the improvement are yet to be determined. Future studies should adapt a 'core battery' of physical function tests to facilitate a detailed comparison of results across exercise modalities. This will enable evidence-based treatment selection according to the defined purpose of training.

Funding/Conflict of interest:

The authors Laurits Taul-Madsen, Luke Connolly, Rachel Dennett, Jenny Freeman, Ulrik Dalgas and Lars G. Hvid declare that they have no conflicts of interest relevant to the content of this article.

References:

1. Compston A, Coles A. Multiple sclerosis. *Lancet (London, England).* 2008;372(9648):1502-1517.
2. Thompson AJ, Baranzini SE, Geurts J, Hemmer B, Ciccarelli O. Multiple sclerosis. *Lancet (London, England).* 2018;391(10130):1622-1636.
3. Calabresi PA. Diagnosis and management of multiple sclerosis. *American family physician.* 2004;70(10):1935-1944.
4. Green R, Cutter G, Friendly M, Kister I. Which symptoms contribute the most to patients' perception of health in multiple sclerosis? *Multiple sclerosis journal - experimental, translational and clinical.* 2017;3(3):2055217317728301.
5. Heesen C, Haase R, Melzig S, et al. Perceptions on the value of bodily functions in multiple sclerosis. *Acta neurologica Scandinavica.* 2018;137(3):356-362.
6. Zhang Y, Taylor BV, Simpson S, Jr., et al. Feelings of depression, pain and walking difficulties have the largest impact on the quality of life of people with multiple sclerosis, irrespective of clinical phenotype. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2020:1352458520958369.
7. Hvid LG, Feys P, Baert I, Kalron A, Dalgas U. Accelerated Trajectories of Walking Capacity Across the Adult Life Span in Persons With Multiple Sclerosis: An Underrecognized Challenge. *Neurorehabilitation and neural repair.* 2020;34(4):360-369.
8. Tremlett H, Paty D, Devonshire V. Disability progression in multiple sclerosis is slower than previously reported. *Neurology.* 2006;66(2):172-177.
9. Kister I, Bacon TE, Chamot E, et al. Natural history of multiple sclerosis symptoms. *International journal of MS care.* 2013;15(3):146-158.
10. Kobelt G, Thompson A, Berg J, Gannedahl M, Eriksson J. New insights into the burden and costs of multiple sclerosis in Europe. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2017;23(8):1123-1136.
11. Ness NH, Schriefer D, Haase R, Ettle B, Cornelissen C, Ziemssen T. Differentiating societal costs of disability worsening in multiple sclerosis. *Journal of neurology.* 2020;267(4):1035-1042.
12. Gyllensten H, Kavaliunas A, Alexanderson K, Hillert J, Tinghög P, Friberg E. Costs and quality of life by disability among people with multiple sclerosis: a register-based study in Sweden. *Multiple sclerosis journal - experimental, translational and clinical.* 2018;4(3):2055217318783352-2055217318783352.
13. Braley TJ, Chervin RD. Fatigue in multiple sclerosis: mechanisms, evaluation, and treatment. *Sleep.* 2010;33(8):1061-1067.
14. Pilutti LA, Platta ME, Motl RW, Latimer-Cheung AE. The safety of exercise training in multiple sclerosis: a systematic review. *Journal of the neurological sciences.* 2014;343(1-2):3-7.
15. Dalgas U, Stenager E, Ingemann-Hansen T. Multiple sclerosis and physical exercise: recommendations for the application of resistance-, endurance- and combined training. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2008;14(1):35-53.
16. Latimer-Cheung AE, Pilutti LA, Hicks AL, et al. Effects of exercise training on fitness, mobility, fatigue, and health-related quality of life among adults with multiple sclerosis: a systematic review to inform guideline development. *Archives of physical medicine and rehabilitation.* 2013;94(9):1800-1828.e1803.
17. Heine M, van de Port I, Rietberg MB, van Wegen EE, Kwakkel G. Exercise therapy for fatigue in multiple sclerosis. *The Cochrane database of systematic reviews.* 2015(9):Cd009956.
18. Snook EM, Motl RW. Effect of exercise training on walking mobility in multiple sclerosis: a meta-analysis. *Neurorehabilitation and neural repair.* 2009;23(2):108-116.

19. Pearson M, Dieberg G, Smart N. Exercise as a therapy for improvement of walking ability in adults with multiple sclerosis: a meta-analysis. *Archives of physical medicine and rehabilitation.* 2015;96(7):1339-1348.e1337.
20. Hobart J, Blight AR, Goodman A, Lynn F, Putzki N. Timed 25-foot walk: direct evidence that improving 20% or greater is clinically meaningful in MS. *Neurology.* 2013;80(16):1509-1517.
21. Baert I, Freeman J, Smedal T, et al. Responsiveness and clinically meaningful improvement, according to disability level, of five walking measures after rehabilitation in multiple sclerosis: a European multicenter study. *Neurorehabilitation and neural repair.* 2014;28(7):621-631.
22. Dalgas U, Langeskov-Christensen M, Stenager E, Riemenschneider M, Hvid LG. Exercise as Medicine in Multiple Sclerosis-Time for a Paradigm Shift: Preventive, Symptomatic, and Disease-Modifying Aspects and Perspectives. *Current neurology and neuroscience reports.* 2019;19(11):88.
23. Dennett R, Madsen LT, Connolly L, Hosking J, Dalgas U, Freeman J. Adherence and drop-out in randomized controlled trials of exercise interventions in people with multiple sclerosis: A systematic review and meta-analyses. *Multiple sclerosis and related disorders.* 2020;43:102169.
24. Collett J, Dawes H, Meaney A, et al. Exercise for multiple sclerosis: a single-blind randomized trial comparing three exercise intensities. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2011;17(5):594-603.
25. Dettmers C, Sulzmann M, Ruchay-Plössl A, Gütler R, Vieten M. Endurance exercise improves walking distance in MS patients with fatigue. *Acta neurologica Scandinavica.* 2009;120(4):251-257.
26. Heine M, Verschuren O, Hoogervorst EL, et al. Does aerobic training alleviate fatigue and improve societal participation in patients with multiple sclerosis? A randomized controlled trial. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2017;23(11):1517-1526.
27. Kjolhede T, Vissing K, de Place L, et al. Neuromuscular adaptations to long-term progressive resistance training translates to improved functional capacity for people with multiple sclerosis and is maintained at follow-up. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2015;21(5):599-611.
28. Dalgas U, Stenager E, Jakobsen J, et al. Resistance training improves muscle strength and functional capacity in multiple sclerosis. *Neurology.* 2009;73(18):1478-1484.
29. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *Bmj.* 2009;339:b2700.
30. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public health reports (Washington, DC : 1974).* 1985;100(2):126-131.
31. Garber CE, Greaney ML, Riebe D, Nigg CR, Burbank PA, Clark PG. Physical and mental health-related correlates of physical function in community dwelling older adults: a cross sectional study. *BMC geriatrics.* 2010;10:6.
32. Kluger BM, Krupp LB, Enoka RM. Fatigue and fatigability in neurologic illnesses: proposal for a unified taxonomy. *Neurology.* 2013;80(4):409-416.
33. Steele J, Androulakis-Korakakis P, Perrin C, et al. Comparisons of Resistance Training and "Cardio" Exercise Modalities as Countermeasures to Microgravity-Induced Physical Deconditioning: New Perspectives and Lessons Learned From Terrestrial Studies. *Front Physiol.* 2019;10:1150.
34. Garber CE, Blissmer B, Deschenes MR, et al. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory,

musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Medicine and science in sports and exercise.* 2011;43(7):1334-1359.

35. Csapo R, Alegre LM. Effects of resistance training with moderate vs heavy loads on muscle mass and strength in the elderly: A meta-analysis. *Scand J Med Sci Sports.* 2016;26(9):995-1006.
36. Schoenfeld BJ, Grgic J, Ogborn D, Krieger JW. Strength and Hypertrophy Adaptations Between Low- vs. High-Load Resistance Training: A Systematic Review and Meta-analysis. *J Strength Cond Res.* 2017;31(12):3508-3523.
37. Smart NA, Waldron M, Ismail H, et al. Validation of a new tool for the assessment of study quality and reporting in exercise training studies: TESTEX. *International journal of evidence-based healthcare.* 2015;13(1):9-18.
38. Suurmond R, van Rhee H, Hak T. Introduction, comparison, and validation of Meta-Essentials: A free and simple tool for meta-analysis. *Research synthesis methods.* 2017;8(4):537-553.
39. Nebiker L, Lichtenstein E, Minghetti A, et al. Moderating Effects of Exercise Duration and Intensity in Neuromuscular vs. Endurance Exercise Interventions for the Treatment of Depression: A Meta-Analytical Review. *Front Psychiatry.* 2018;9:305.
40. Harvey L, Smith A, Jones R. The Effect of Weighted Leg Raises on Quadriceps Strength, EMG Parameters and Functional Activities in People with Multiple Sclerosis. *Physiotherapy.* 1999;85:154-161.
41. Kinney AR, Eakman AM, Graham JE. Novel Effect Size Interpretation Guidelines and an Evaluation of Statistical Power in Rehabilitation Research. *Archives of physical medicine and rehabilitation.* 2020;101(12):2219-2226.
42. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Bmj.* 2003;327(7414):557-560.
43. Ahmadi A, Arastoo AA, Nikbakht M. The effects of a treadmill training programme on balance, speed and endurance walking, fatigue and quality of life in people with multiple scler1osis : original research. *International SportMed Journal.* 2010;11(4):389-397.
44. Hebert JR, Corboy JR, Manago MM, Schenkman M. Effects of vestibular rehabilitation on multiple sclerosis-related fatigue and upright postural control: a randomized controlled trial. *Phys Ther.* 2011;91(8):1166-1183.
45. Mokhtarzade M, Ranjbar R, Majdinasab N, Patel D, Molanouri Shamsi M. Effect of aerobic interval training on serum IL-10, TNFα, and adipokines levels in women with multiple sclerosis: possible relations with fatigue and quality of life. *Endocrine.* 2017;57(2):262-271.
46. Petajan JH, Gappmaier E, White AT, Spencer MK, Mino L, Hicks RW. Impact of aerobic training on fitness and quality of life in multiple sclerosis. *Annals of Neurology.* 1996;39(4):432-441.
47. Schulz KH, Gold SM, Witte J, et al. Impact of aerobic training on immune-endocrine parameters, neurotrophic factors, quality of life and coordinative function in multiple sclerosis. *Journal of the neurological sciences.* 2004;225(1-2):11-18.
48. Langeskov-Christensen M, Grøndahl Hvid L, Nygaard MKE, et al. Efficacy of high-intensity aerobic exercise on brain MRI measures in multiple sclerosis. *Neurology.* 2020.
49. Tollár J, Nagy F, Tóth BE, et al. Exercise Effects on Multiple Sclerosis Quality of Life and Clinical-Motor Symptoms. *Medicine and science in sports and exercise.* 2020;52(5):1007-1014.
50. Baquet L, Hasselmann H, Patra S, et al. Short-term interval aerobic exercise training does not improve memory functioning in relapsing-remitting multiple sclerosis-a randomized controlled trial. *PeerJ.* 2018;6:e6037.
51. Sadeghi Bahmani D, Razazian N, Farnia V, Alikhani M, Tatari F, Brand S. Compared to an active control condition, in persons with multiple sclerosis two different types of exercise training improved sleep and depression, but not fatigue, paresthesia, and intolerance of uncertainty. *Multiple sclerosis and related disorders.* 2019;36:101356.

52. Feys P, Moumdjian L, Van Halewyck F, et al. Effects of an individual 12-week community-located "start-to-run" program on physical capacity, walking, fatigue, cognitive function, brain volumes, and structures in persons with multiple sclerosis. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2019;25(1):92-103.
53. Mostert S, Kesselring J. Effects of a short-term exercise training program on aerobic fitness, fatigue, health perception and activity level of subjects with multiple sclerosis. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2002;8(2):161-168.
54. Oken BS, Kishiyama S, Zajdel D, et al. Randomized controlled trial of yoga and exercise in multiple sclerosis. *Neurology.* 2004;62(11):2058-2064.
55. Dodd KJ, Taylor NF, Shields N, Prasad D, McDonald E, Gillon A. Progressive resistance training did not improve walking but can improve muscle performance, quality of life and fatigue in adults with multiple sclerosis: a randomized controlled trial. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2011;17(11):1362-1374.
56. Moradi M, Sahraian MA, Aghsaie A, et al. Effects of Eight-week Resistance Training Program in Men With Multiple Sclerosis. *Asian J Sports Med.* 2015;6(2):e22838.
57. Callesen J, Cattaneo D, Brincks J, Kjeldgaard Jørgensen M-L, Dalgas U. How do resistance training and balance and motor control training affect gait performance and fatigue impact in people with multiple sclerosis? A randomized controlled multi-center study. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2019:1352458519865740-1352458519865740.
58. Kjølhede T, Vissing K, de Place L, et al. Neuromuscular adaptations to long-term progressive resistance training translates to improved functional capacity for people with multiple sclerosis and is maintained at follow-up. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2015;21(5):599-611.
59. Hosseini SS, Rajabi H, Sahraian MA, Moradi M, Mehri K, Abolhasani M. Effects of 8-Week Home-Based Yoga and Resistance Training on Muscle Strength, Functional Capacity and Balance in Patients with Multiple Sclerosis: A Randomized Controlled Study. *Asian journal of sports medicine.* 2018;9.
60. DeBolt LS, McCubbin JA. The effects of home-based resistance exercise on balance, power, and mobility in adults with multiple sclerosis. *Archives of physical medicine and rehabilitation.* 2004;85(2):290-297.
61. Dalgas U, Stenager E, Jakobsen J, et al. Fatigue, mood and quality of life improve in MS patients after progressive resistance training. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2010;16(4):480-490.
62. Langeskov-Christensen M, Heine M, Kwakkel G, Dalgas U. Aerobic capacity in persons with multiple sclerosis: a systematic review and meta-analysis. *Sports medicine (Auckland, NZ).* 2015;45(6):905-923.
63. Jørgensen M, Dalgas U, Wens I, Hvid LG. Muscle strength and power in persons with multiple sclerosis - A systematic review and meta-analysis. *Journal of the neurological sciences.* 2017;376:225-241.
64. Kjølhede T, Vissing K, Dalgas U. Multiple sclerosis and progressive resistance training: a systematic review. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2012;18(9):1215-1228.
65. Langeskov-Christensen D, Feys P, Baert I, Riemenschneider M, Stenager E, Dalgas U. Performed and perceived walking ability in relation to the Expanded Disability Status Scale in persons with multiple sclerosis. *Journal of the neurological sciences.* 2017;382:131-136.
66. Nieuwenhuis MM, Van Tongeren H, Sørensen PS, Ravnborg M. The six spot step test: a new measurement for walking ability in multiple sclerosis. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2006;12(4):495-500.

67. Sieljacks PS, Søberg CA, Michelsen AS, Dalgas U, Hvid LG. Lower extremity muscle strength across the adult lifespan in multiple sclerosis: Implications for walking and stair climbing capacity. *Experimental gerontology.* 2020;139:111025.
68. Moss-Morris R, Harrison AM, Safari R, et al. Which behavioural and exercise interventions targeting fatigue show the most promise in multiple sclerosis? A systematic review with narrative synthesis and meta-analysis. *Behaviour research and therapy.* 2019:103464.
69. Sabapathy NM, Minahan CL, Turner GT, Broadley SA. Comparing endurance- and resistance-exercise training in people with multiple sclerosis: a randomized pilot study. *Clinical rehabilitation.* 2011;25(1):14-24.
70. Gijbels D, Dalgas U, Romberg A, et al. Which walking capacity tests to use in multiple sclerosis? A multicentre study providing the basis for a core set. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2012;18(3):364-371.
71. Andreasen AK, Stenager E, Dalgas U. The effect of exercise therapy on fatigue in multiple sclerosis. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2011;17(9):1041-1054.
72. Rooney S, Wood L, Moffat F, Paul L. Is Fatigue Associated With Aerobic Capacity and Muscle Strength in People With Multiple Sclerosis: A Systematic Review and Meta-analysis. *Archives of physical medicine and rehabilitation.* 2019;100(11):2193-2204.
73. Paul L, Coote S, Crosbie J, et al. Core outcome measures for exercise studies in people with multiple sclerosis: recommendations from a multidisciplinary consensus meeting. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2014;20(12):1641-1650.
74. Edwards T, Pilutti LA. The effect of exercise training in adults with multiple sclerosis with severe mobility disability: A systematic review and future research directions. *Multiple sclerosis and related disorders.* 2017;16:31-39.

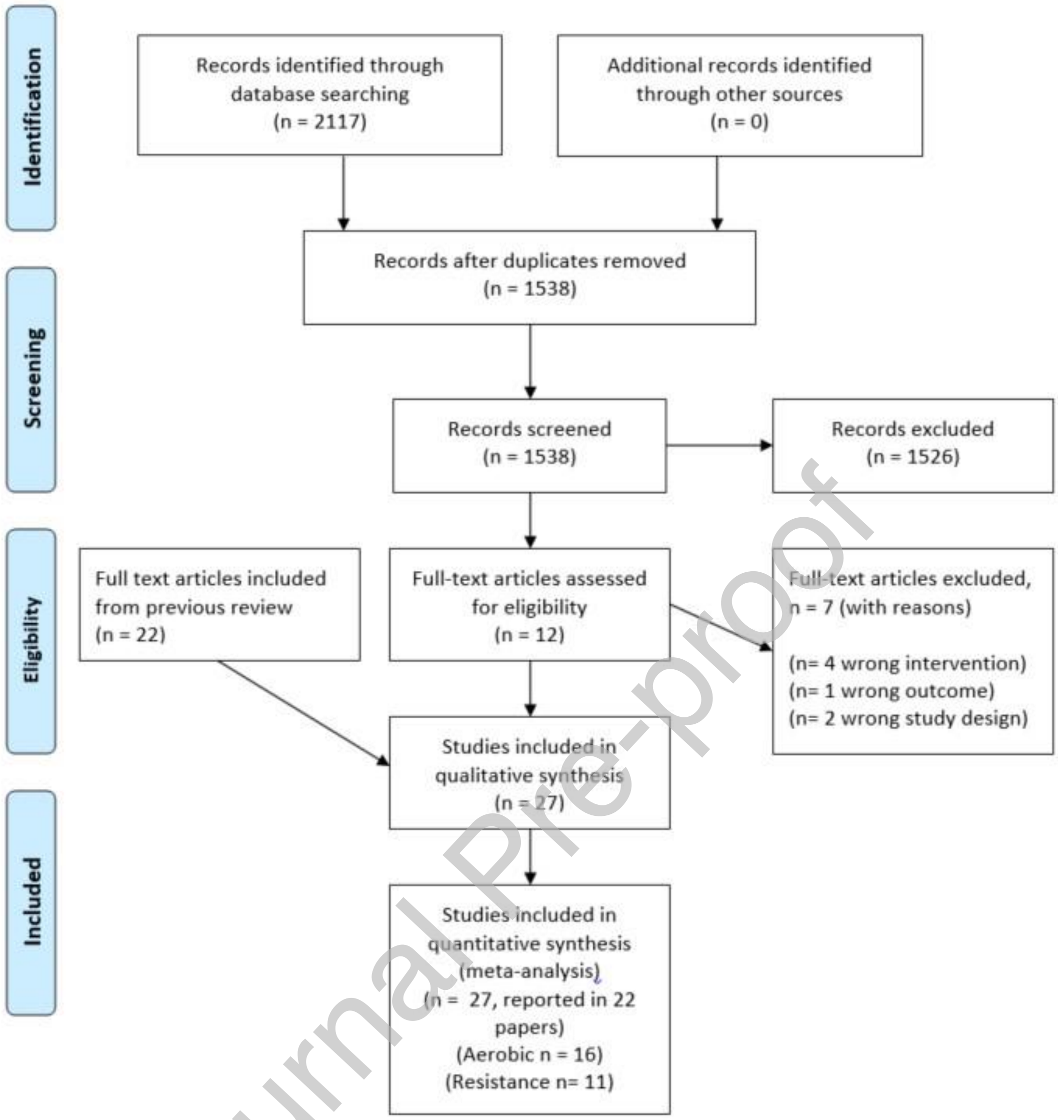
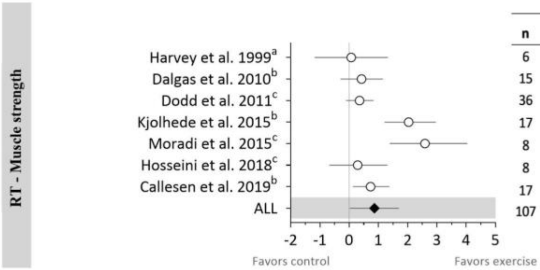
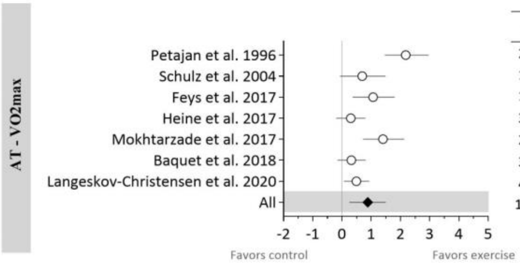


Figure 1: PRISMA flow diagram on the search result and study selection process.



Study	Exercise (RT) n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Control n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Weight	SMD (random), 95% CI
Harvey et al. 1999[a]	6	17.8	4.0	24.2	3.1	5	21.1	13.3	23.6	11.4	11.8%	0.07 [-1.18; 1.32]
Dalgas et al. 2010[b]	15	124	51	138	47	16	118	38	119	43	15.6%	0.42 [-0.30; 1.15]
Dodd et al. 2011[c]	36	70	36	86	66	35	62	38	66	42	17.9%	0.35 [-0.11; 0.83]
Kjolhede et al. 2015[b]	17	1.17	0.10	1.44	0.10	15	1.16	0.11	1.22	0.11	14.1%	2.03 [1.21; 2.96]
Moradi et al. 2015[c]	8	115	26	147	23	10	100	23	88	21	10.6%	2.59 [1.39; 4.04]
Hosseini et al. 2018[c]	8	126	56	148	46	8	128	56	132	60	13.5%	0.29 [-0.68; 1.30]
Callesen et al. 2019[b]	17	1.27	0.39	1.51	0.36	18	1.17	0.49	1.23	0.41	16.5%	0.73 [0.12; 1.37]
ALL	107					107					100.0%	**0.86 [0.02; 1.70]**

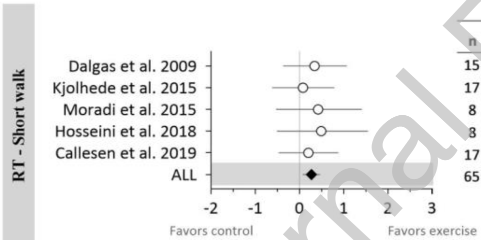
Overall effect: $p = 0.013$. Heterogeneity: $I^2 = 75\%$.



Study	Exercise (AT) n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Control n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Weight	SMD (random), 95% CI
Petajan et al. 1996	21	24.2	1.4	29.4	1.3	25	26.0	1.3	26.4	1.4	11.8%	2.18 [1.47; 2.96]
Schulz et al. 2004	15	33.1	7.1	36.5	7.6	13	28.9	7.8	31.3	6.9	15.6%	0.69 [-0.07; 1.49]
Feys et al. 2017	18	23.9	5.9	25.4	5.0	17	21.8	4.0	20.1	4.8	17.9%	1.06 [0.36; 1.80]
Heine et al. 2017	33	22.6	6.2	23.5	7.4	30	20.8	5.6	21.4	6.4	14.1%	0.30 [-0.20; 0.80]
Mokhtarzade et al. 2017	22	17.8	1.6	20.0	1.4	18	16.7	1.4	17.0	2.7	10.6%	1.40 [0.72; 2.13]
Baquet et al. 2018	34	27.2	7.9	28.0	8.8	34	25.6	5.5	25.6	5.4	13.5%	0.32 [-0.15; 0.81]
Langeskov-Christensen et al. 2020	43	28.2	6.9	32.1	7.3	43	28.6	7.7	28.4	7.6	16.5%	0.49 [0.07; 0.93]
All	186					180					100.0%	**0.88 [0.25; 1.50]**

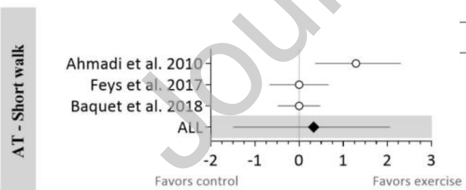
Overall effect: $p = 0.001$. Heterogeneity: $I^2 = 78\%$.

Figure 2: Meta-analysis of the effect of aerobic training and resistance training on physiological adaptations. Abbreviations: RT: Resistance training; AT: Aerobic training: VO_{2max}: Maximal oxygen consumption a: Strength measured in knee extensor; b: Strength measured in knee extensor and flexor (average); c: Strength measured in legpress



Study	Exercise (RT) n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Control n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Weight	SMD (random), 95% CI
Dalgas et al. 2009	15	1.30	0.68	1.52	0.79	16	1.37	0.52	1.27	0.64	23.2%	0.34 [-0.37; 1.07]
Kjolhede et al. 2015	17	1.66	0.28	1.83	0.28	15	1.80	0.34	1.80	0.41	24.2%	0.08 [-0.62; 0.79]
Moradi et al. 2015	8	1.20	0.65	1.52	0.46	10	1.37	0.30	1.35	0.29	13.8%	0.42 [-0.53; 1.41]
Hosseini et al. 2018	8	1.47	0.28	1.49	0.25	8	1.37	0.30	1.35	0.29	12.5%	0.49 [-0.51; 1.55]
Callesen et al. 2019	17	1.24	0.34	1.32	0.43	18	1.22	0.40	1.23	0.45	26.3%	0.20 [-0.47; 0.88]
ALL	65					67					100.0%	**0.27 [0.07; 0.47]**

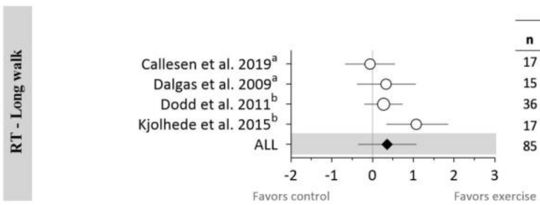
Overall effect: $p = 0.000$. Heterogeneity: $I^2 = 0\%$.



Study	Exercise (AT) n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Control n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Weight	SMD (random), 95% CI
Ahmadi et al. 2010	11	1.16	0.26	1.43	0.32	10	1.09	0.22	1.06	0.22	25.7%	1.29 [0.37; 2.31]
Feys et al. 2017	18	1.86	0.23	1.91	0.23	17	1.91	0.29	1.91	0.29	34.2%	0.00 [-0.67; 0.67]
Baquet et al. 2018	34	1.59	0.26	1.59	0.26	34	1.59	0.26	1.59	0.26	40.1%	0.00 [-0.48; 0.48]
ALL	63					61					100.0%	**0.33 [-1.49; 2.06]**

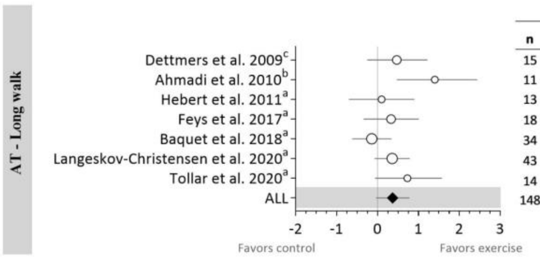
Overall effect: $p = 0.202$. Heterogeneity: $I^2 = 69\%$.

Figure 3: Meta-analysis of the effect of aerobic training and resistance training on the performance of a short walking test. Abbreviations: RT: Resistance training; AT: Aerobic training



Study	Exercise (RT) n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Control n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Weight	SMD (random), 95% CI
Callesen et al. 2019[a]	17	380	124	392	144	18	398	122	401	146	25.8%	-0.06 [-0.67; 0.55]
Dalgas et al. 2009[a]	15	441	187	495	186	16	438	143	436	164	21.7%	0.33 [-0.38; 1.06]
Dodd et al. 2011[b]	36	120	36	123	35	35	112	37	113	39	32.1%	0.27 [-0.20; 0.74]
Kjolhede et al. 2015[b]	17	192	12	214	12	15	196	12	200	12	20.4%	1.07 [0.34; 1.86]
ALL	85					84					100.0%	**0.36 [-0.35; 1.08]**

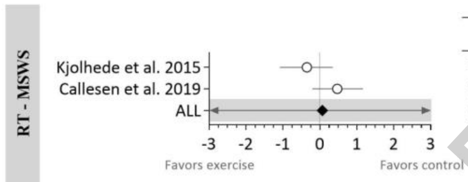
Overall effect: p = 0.108. Heterogeneity: I^2 = 48%.



Study	Exercise (AT) n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Control n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Weight	SMD (random), 95% CI
Dettmers et al. 2009[c]	15	1043	564	1693	978	15	1163	750	1260	794	12.8%	0.47 [-0.25; 1.22]
Ahmadi et al. 2010[b]	11	122	21	142	21	10	122	28	110	23	8.8%	1.40 [0.47; 2.44]
Hebert et al. 2011[a]	13	325	102	339	119	12	320	100	327	114	11.6%	0.10 [-0.70; 0.90]
Feys et al. 2017[a]	18	576	61	590	50	17	574	67	570	69	14.1%	0.33 [-0.34; 1.01]
Baquet et al. 2018[a]	34	432	106	454	104	34	449	80	467	84	19.8%	-0.14 [-0.62; 0.34]
Langeskov-Christensen et al. 2020[a]	43	531	103	579	97	43	521	129	537	133	21.6%	0.36 [-0.07; 0.79]
Tollar et al. 2020[a]	14	246	41	278	39	12	243	40	250	35	11.4%	0.73 [-0.06; 1.57]
ALL	148					143					100.0%	**0.37 [-0.04; 0.78]**

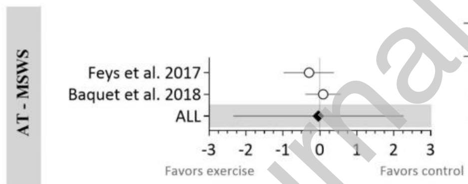
Overall effect: p = 0.026. Heterogeneity: I^2 = 43%.

Figure 4: Meta-analysis of the effect of aerobic training and resistance training on the performance of a long walking test. Abbreviations: RT: Resistance training; AT: Aerobic training a: 6 minute walk test, b: 2 minute walk test, c: Maximum walking distance



Study	Exercise (RT) n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Control n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Weight	SMD (random), 95% CI
Kjolhede et al. 2015	17	32.50	5.30	24.00	5.40	15	27.20	5.50	26.00	5.60	49.4%	-0.35 [-1.08; 0.35]
Callesen et al. 2019	17	44.30	10.40	39.40	10.60	18	36.00	9.20	34.20	10.80	50.6%	0.47 [-0.20; 1.17]
ALL	34					33					100.0%	**0.07 [-5.20; 5.33]**

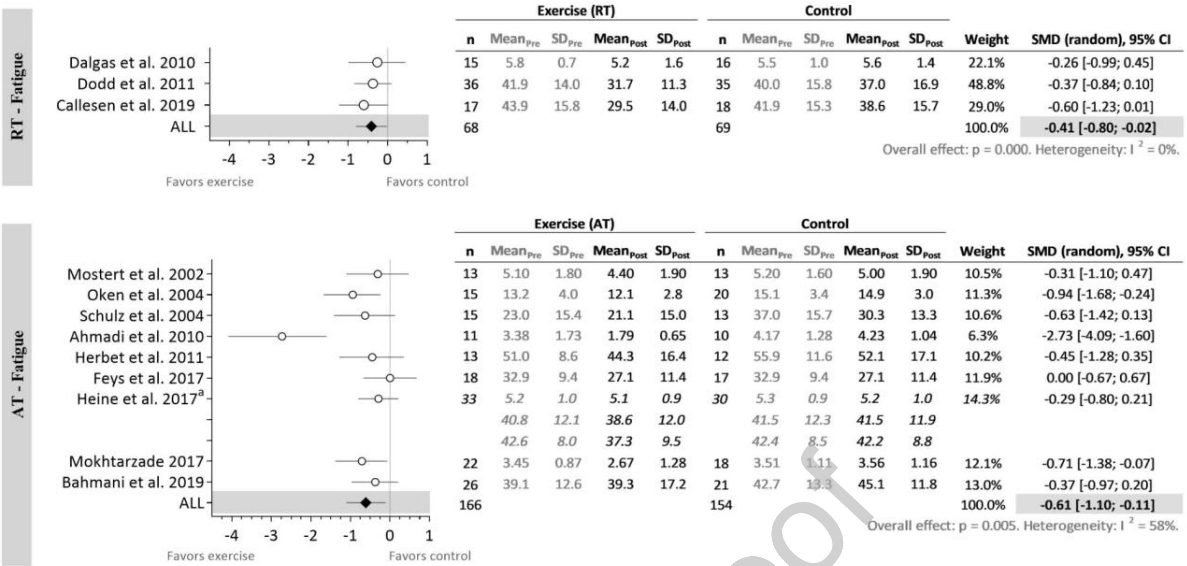
Overall effect: p = 0.875. Heterogeneity: I^2 = 66%.



Study	Exercise (AT) n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Control n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Weight	SMD (random), 95% CI
Feys et al. 2017	18	19.10	16.40	15.00	12.80	17	16.30	18.90	21.10	26.10	34.3%	-0.29[-0.98; 0.38]
Baquet et al. 2018	34	19.40	10.40	19.60	10.20	34	18.70	10.70	18.70	9.90	65.7%	0.09 [-0.39; 0.57]
ALL	52					51					100.0%	**-0.04 [-2.34; 2.26]**

Overall effect: p = 0.816. Heterogeneity: I^2 = 0%.

Figure 5: Meta-analysis of the effect of aerobic training and resistance training on self-reported walking ability. Abbreviations: RT: Resistance training; AT: Aerobic training



RT - Fatigue

Study	Exercise (RT) n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Control n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Weight	SMD (random), 95% CI
Dalgas et al. 2010	15	5.8	0.7	5.2	1.6	16	5.5	1.0	5.6	1.4	22.1%	-0.26 [-0.99; 0.45]
Dodd et al. 2011	36	41.9	14.0	31.7	11.3	35	40.0	15.8	37.0	16.9	48.8%	-0.37 [-0.84; 0.10]
Callesen et al. 2019	17	43.9	15.8	29.5	14.0	18	41.9	15.3	38.6	15.7	29.0%	-0.60 [-1.23; 0.01]
ALL	68					69					100.0%	**-0.41 [-0.80; -0.02]**

Overall effect: p = 0.000. Heterogeneity: I^2 = 0%.

AT - Fatigue

Study	Exercise (AT) n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Control n	$Mean_{Pre}$	SD_{Pre}	$Mean_{Post}$	SD_{Post}	Weight	SMD (random), 95% CI
Mostert et al. 2002	13	5.10	1.80	4.40	1.90	13	5.20	1.60	5.00	1.90	10.5%	-0.31 [-1.10; 0.47]
Oken et al. 2004	15	13.2	4.0	12.1	2.8	20	15.1	3.4	14.9	3.0	11.3%	-0.94 [-1.68; -0.24]
Schulz et al. 2004	15	23.0	15.4	21.1	15.0	13	37.0	15.7	30.3	13.3	10.6%	-0.63 [-1.42; 0.13]
Ahmadi et al. 2010	11	3.38	1.73	1.79	0.65	10	4.17	1.28	4.23	1.04	6.3%	-2.73 [-4.09; -1.60]
Herbet et al. 2011	13	51.0	8.6	44.3	16.4	12	55.9	11.6	52.1	17.1	10.2%	-0.45 [-1.28; 0.35]
Feys et al. 2017	18	32.9	9.4	27.1	11.4	17	32.9	9.4	27.1	11.4	11.9%	0.00 [-0.67; 0.67]
Heine et al. 2017[a]	*33*	*5.2*	*1.0*	*5.1*	*0.9*	*30*	*5.3*	*0.9*	*5.2*	*1.0*	*14.3%*	-0.29 [-0.80; 0.21]
		40.8	*12.1*	*38.6*	*12.0*		*41.5*	*12.3*	*41.5*	*11.9*		
		42.6	*8.0*	*37.3*	*9.5*		*42.4*	*8.5*	*42.2*	*8.8*		
Mokhtarzade 2017	22	3.45	0.87	2.67	1.28	18	3.51	1.11	3.56	1.16	12.1%	-0.71 [-1.38; -0.07]
Bahmani et al. 2019	26	39.1	12.6	39.3	17.2	21	42.7	13.3	45.1	11.8	13.0%	-0.37 [-0.97; 0.20]
ALL	166					154					100.0%	**-0.61 [-1.10; -0.11]**

Overall effect: p = 0.005. Heterogeneity: I^2 = 58%.

Figure 6: Meta-analysis of the effect of aerobic training and resistance training on the perceptions of fatigue. Abbreviations: RT: Resistance training; AT: Aerobic training a: Effect size as an average of the Fatigue Severity Scale, Modified Fatigue Impact Scale and the CIS20r: Checklist Individual Strength

Table 1: Characteristics of included studies

Study	Active/ passive control	Recruited (n)	Analysed (n)	Sex (% female)	Age (Years mean)	MS type (% RRMS)	EDSS (0-10)	Disease duration (mean)	Individual/ group	Supervised/ unsupervised	Frequency (d/wk) and duration (wks)	Intensity	Length of training session (min)
Aerobic training													
Ahmadi, 2010[55] 2013[72] and Arastoo 2013[73]	Waitlist control	Int: 11 Con :10	NR NR	100 100	32.2 36.7	NR NR	2.0 2.3	4.7 5.0	NR N/A	Supervised N/A	3d/wk *8wk NR	40-75 % of HR_{max} N/A	30 N/A
Bahmani, 2019[56]	Active control (attention control)	Int: 31 Con:31	26 21	100 100	38.0 37.9	NR	2.5 2.0	6.9 7.2	Group Group	Supervised Supervised	3d/wk *8wk 3d/wk *8wk	NR N/A	30-45 30-45
Baquet, 2018[41]	Waitlist control (usual activity)	Int: 34 Con:34	34 34	62 74	38.2 39.6	100 100	1.7 1.8	6.8 5.7	Group N/A	Supervised N/A	2.5d/ wk*12wk N/A	RPE of 8 NR	27-69 N/A
Dettmers, 2009[25]	Active control (stretching and relaxation)	Int: 15 Con:15	15 15	67 73	45.8 39.7	77	2.6 2.8	10.7 10.5	Group Group	Supervised Supervised	3d/wk *3wk 3d/wk *3wk	NR NR	45 45
Feys, 2017[44]	Waitlist control (usual activity)	Int: 21 Con: 21	18 17	95 86	36.6 44.4	NR	NR	8.1 9.2	Indiv N/A	Unsupervised N/A	3d/wk *12wk N/A	NR N/A	NR N/A
Heine, 2017[26]	Active control (consultation with MS nurse)	Int1: 43 Con: 46	33 30	74 72	48.8 48.4	79 74	2.5 3.0	7.0 12.0	Group NR	Unsupervised NR	3d/wk *16wks NR	3min 40%, 1 min 60% and 1min at 80% of PPO. N/A	30 NR
Hebert, 2011[45]	Waitlist control (usual activity)	Int: 13 Con: 13	13 12	85 85	42.6 50.	85 92	NR NR	5.1 9.1	NR N/A	Supervised N/A	2d/wk *6wk N/A	65-75% of HR_{max} N/A	60 N/A

					2									
Mokhtarzade, 2017[49]	Passive control	Int : 25 Con:20	2218	10100	32.031.3	10100	1.841.57	2.693.47	NR N/A	NR N/A	3d/wk *8wk N/A	60-75% of W_{max} N/A	42-66 N/A	
Langeskov-Christensen, 2020[46]	Waitlist control (habitual activity)	Int : 43 Con:43	4343	6060	44.045.6	9579	2.72.8	10.98.6	Group N/A	Supervised N/A	2d/wk *24wk N/A	65-95% of HR_{max} N/A	30-60 N/A	
Mostert, 2002[57]	Active control (usual activity)	Int: 18 Con:18	1313	7785	45.243.9	30.838.5	4.64.5	11.211.6	NR N/A	Supervised N/A	5d/wk *4wk N/A	NR N/A	30 N/A	
Oken, 2004[47]	Waitlist control (usual activity)	Int: 21 Con: 22	1520	87100	48.848.4	NR	2.93.1	NR	Indiv N/A	Supervised N/A	1d/wk *26wk N/A	NR N/A	NR N/A	
Petajan, 1996[50]	Waitlist control (usual activity)	Int: 21 Con:25	2125	7164	41.139.0	NR	3.82.9	9.36.2	NR N/A	Supervised N/A	3d/wk *15wk N/A	60% of VO_{2max} N/A	50 N/A	
Schulz, 2004[51]	Waitlist control (usual activity)	Int::15 Con:13	1513	7362	39.042.0	NRNR	2.02.5	NRNR	NR NR	NR NR	2d/wk *8wk N/A	75% of W_{max} NR	30 N/A	
Tollar, 2020[48]	Waitlist control (usual activity)	Int: 14 Con:12	1412	9392	48.144.4	5066	5-65-6	13.214.0	Group N/A	Supervised Supervised	5d/wk *5wk N/A	80% of age-predicted HR_{max} N/A	60 N/A	
Progressive resistance training														
Callesen, 2019[42]	Passive control (usual activity)	Int: 23 Con:20	1718	7080	52.056.	7065	4.03.5	15.011.	Group N/A	Supervised N/A	2d/wk *10wk N/A	10 sets at 15 RM – 8 sets at 8 RM N/A	NR N/A	

					0			0					
Dalgas, 2009[28] and 2010[58,74]	Passive control (waitlist usual activity)	Int: 19 Con: 19	1516	6662	47.750.4	100	3.73.9	6.68.1	Group N/A	Supervised N/A	2d/wk *12wk N/A	3-4 sets of 8-12 repetitions at 8-15 RM N/A	NR N/A
DeBolt, 2004[40]	Passive control (usual activity)	Int: 19 Con: 18	1917	7978	51.647.8	4744	4.03.5	15.113.0	Indiv N/A	Unsupervised N/A	3d/wk *8wk N/A	2-3 sets of 8-12 repetitions wearing a weighted vest (0.5% of BW) increasing by 0.5-1.5% of BW every 2 wk	35-50 N/A
Dodd, 2011[43]	Passive control (Usual activity + social program)	Int: 39 Con: 37	3635	7274	47.750.4	100	NR	NR	Group Group	Supervised Supervised	2d/wk *10wk 1d/wk *10wk	2 sets of 10-12 repetitions at 10-12 RM N/A	4560
Harvey, 1999[36]	Passive control (usual activity)	Int : 7 Con: 5	65	8380	38.043.0	100	NR	510	Indiv N/A	Unsupervised N/A	2d/*8wk N/A	5 sets of 10 leg extensions using 0.5 or 1kg ankle weights N/A	NR N/A
Hosseini, 2018[52]	Passive control (usual activity)	Int: 9 Con: 8	88	5550	32.933.0	NR	NR	NR	Indiv N/A	Unsupervised N/A	3d/wk *8wk N/A	1% of BW fastened to body increasing by 0.5-1% every 2 wk N/A	35-50 N/A
Kjolhede, 2015[53] and Jørgensen 2019[75]	Passive control (waitlist usual activity)	Int: 18 Con: 17	1715	NR	43.2	100	3	5	NR N/A	Supervised N/A	2d/wk *24wk N/A	3-5 sets of 6-10 repetitions at 6-15 RM N/A	NR N/A
Moradi, 2015[54]	Passive control (usual activity)	Int: 10 Con: 10	810	00	34.433.1	6260	3.03.0	8.116.5	NR N/A	Supervised N/A	3d/wk *8wk N/A	1 set of 6-15 repetitions at 50-80% of 1 RM N/A	30 N/A

Abbreviations: MS: Multiple sclerosis; RRMS: Relapse remitting multiple sclerosis; Int: Intervention; Con: Control; PPO: Peak power output achieved during incremental exercise test to exhaustion; VO_{2max}: Maximal oxygen consumption; VO_{2peak}: Peak oxygen consumption; HR: Heart rate; RM: Repetition maximum; RPE: ratings of perceived exertion; BW: Body weight; MVC: Maximal voluntary contraction;

MIP: Maximal inspiratory pressure; TMW: Tolerated maximum workload; W: Watts; Indiv: Individual; N/A: not applicable; NR: not reported.

Table 2: TESTEX scores

Paper	Eligibility criteria	Randomi-zation	Allocation concealed	Baseline data	assessor primary OM	OM in >85% patients	AE reported	Exercise attendance	Intention-to-treat	stats primary OM	stats secondary OM	Outcomes point estimates	Control Physical activity	Exercise load titrated	calculated	**Total**
Ahmadi 2010	1	0	1	1	0	1	0	0	0	0	1	1	0	1	1	**8**
Ahmadi 2013a	1	0	1	1	0	1	0	0	0	0	1	1	0	0	0	**6**
Ahmadi 2013b	1	0	1	1	0	1	0	0	0	0	1	1	0	1	1	**8**
Bahmani 2019	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	**6**
Baquet 2018	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	**13**
Callesen 2019	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	**12**
Dalgas 2009	1	1	1	1	1	0	0	1	0	1	1	1	0	1	1	**11**
Dalgas 2010a	1	0	1	1	0	0	0	1	0	1	1	1	0	1	1	**9**
Dalgas 2010b	1	0	1	1	1	0	0	1	0	1	1	1	0	1	1	**10**
DeBolt 2004	1	0	0	1	0	1	0	1	0	1	1	1	0	1	1	**9**
Dettmers 2009	1	1	0	1	0	1	0	0	0	1	1	0	0	0	0	**6**
Dodd 2011	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	**13**
Feys 2017	1	0	0	1	0	0	1	1	1	1	1	1	0	0	1	**9**
Harvey 1999	1	1	0	1	0	1	0	1	0	1	1	0	0	0	0	**7**
Hebert 2011	1	0	1	1	1	0	1	1	1	1	1	1	0	0	1	**11**
Heine 2017	1	1	0	0	1	0	0	1	1	1	1	1	1	1	1	**11**
Hosseini 2018	1	1	0	1	0	1	0	0	0	0	0	1	0	0	0	**5**
Jørgensen 2019	1	0	1	1	0	1	0	0	0	0	1	0	0	1	1	**7**
Kjølhede 2015	1	1	1	0	0	1	0	1	1	1	1	1	0	1	1	**11**
Langeskov-Christensen 2020	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	**13**
Mokhtarzade 2017	1	0	1	1	0	1	0	0	0	0	0	1	0	1	1	**7**
Moradi 2015	1	1	0	1	1	1	1	0	0	1	1	1	0	1	1	**11**
Mostert 2002	1	0	1	1	0	0	0	1	0	0	0	0	0	0	0	**4**
Oken 2004	1	0	1	1	1	0	1	1	0	1	0	1	0	0	0	**8**
Petajan 1996	1	0	0	1	1	1	0	1	0	0	0	1	0	1	1	**8**
Schulz 2004	1	0	1	1	0	0	0	0	0	1	1	1	0	0	1	**7**
Tollar 2020	1	1	1	1	1	1	1	1	0	1	1	1	0	0	0	**11**

Abbreviations: OM: Outcome measure

Table 3: Effect sizes of all outcomes.

Study	Strength/ VO2peak (positive ES = improvement)		Short walk (Positive ES = improvement)			Long walk (Positive ES = improvement)			Other walking (Negative ES = improvement)	Functional mobility (other) (Negative ES = improvement)				Perceived fatigue (Negative ES = improvement)			
	Strength	VO_{2peak}	T25FW	T10MW	T50MW	2minW	6minW	Distance	MSWS	TUG	SSST	5-STS	Stair climb	FSS	MFIS	CIS20r	FSMC
Aerobic training																	
Ahmadi, Arastoo 2010+13				**1.29**		**1.40**								**-2.73**			
Dettmers 2009								**0.47**									
Bahmani 2019														-0.37			
Baquet 2018		0.32	0.00				-0.14		0.09								
Feys 2017		**1.06**	0.00				0.33		**-0.29**			**-0.38**					0.0
Heine 2017		0.30												-0.29	-0.29	**-0.52**	
Hebert 2011							0.10								**-0.45**		
Langeskov-Christensen 2020		**0.49**					0.36										
Mokhtarzade 2017		**1.40**												**-0.71**			
Mostert 2002														-0.31			
Oken 2004															-0.94		
Petajan 1996		**2.18**															
Schulz 2004		0.69													-0.63		
Tollar 2020							**0.73**										
Resistance training																	
Callesen 2019	**0.73**		0.20				1.07		0.47		-0.18				**-0.60**		
Dalgas 2009+10	**0.42**			0.34			-0.06					-0.83	-0.52	**-0.26**			
DeBolt 2004										-0.35							
Dodd 2011	**0.35**					0.27									**-0.37**		
Hosseini 2018	0.29			0.49													
Harvey 1999*	0.07			n.c	n.c.												
Kjølhede 2015, Jørgensen 2019	**2.03**		**0.08**			**0.27**			**-0.35**			**-2.27**	**-1.84**				
Moradi 2015	**2.59**			0.42						**-0.**							

				87	

Abbreviations: ES: Effect size; $VO2_{peak}$: Peak oxygen consumption; T25FW: Timed 25 foot walk; T10MW: Timed 10 meter walk; T50MW: Timed 50 meter foot walk; TUG: Timed up and go; SSST: Six spot step test; 5-STS: 5 times sit to stand; FSS: Fatigue Severity Scale; MFIS: Modified Fatigue Impact Scale; CIS20r: Checklist Individual Strength. *ES was non computable as no standard deviation was reported.
Bold *indicates that the paper has reported a statistically significant between-group change.*